国語教育学

国語の学力測定

田中寛一

丸山良二

國語科學講座

─XI─

國語教育學

國語の學力測定

田中寛一・丸山良二

株式會社
明治書院

國語科學講座

— XI —

國語教育學

國語の學力測定

田中寛一・丸山良二

株式會社

明治書院

目次

國語の學力測定

田中寛一
丸山良二

一 學力測定の性質

兒童は生まれながらに自己活動的で、外界からの刺戟を求め、その刺戟を同化し組織し、以て自己發展を繼續して居る。學校敎師の敎授上の任務はかゝる兒童に對して價値ある敎材を提供して、これを攝取せしめ組織せしめ、更に進んで價値ある知識技能を學習しようとする態度を養成し、以て兒童の正しい方向への自己發展を助成するにある。

かくて一度敎師が自己の敎授方法に就いて反省するならば、少くとも敎師は、

(一)敎授が成功して居るか否かを知るために何等かの方法を用ひて兒童に生じた組織發展の蹟を調べてみなければならぬ。

(二)若し兒童にして豫期の成績を修めて居ないものが居るならばその原因理由を診斷し發見しなければならぬ。

(三)成績劣等の原因が分つたならばこれに對する對策を講じ、敎師の手に於て治療し得るものならば出來る限り早

期に治療してやることが肝要である。

（四）兒童をして學力進歩の蹟を知らせてその努力を刺戟しなければならぬ。

上記の如き目的を達するために用ひる方法を試驗又は檢査といふ。試驗又は檢査の結果を整理して一定兒童が多人數の中に於て占める位置を決定することを考査といふ。試驗又は考査は敎師が口頭で尋ねた際に兒童も口答するが如き極めて略式のものもあれば、また兒童をして多數の問題に筆答させるが如き正式のものもある。略式・正式の何れにせよ、試驗は敎育的手段によつて兒童に生じた進歩を測定し、その結果によつて一は敎師の敎授方法反省の資料とし、他は兒童の學習努力を刺戟しようとするものである。

考査の目的が以上の如きものだけであるとしても、その材料を得る方法としての試驗方法は正確で信賴し得るものでなければならぬことは明かである。況して學級相互の成績を比較するとか、又は學校相互の成績を比較するとか、或は新しい敎授法の敎育的有效さを決定するとか、兒童の進級又は卒業を決定するが如き場合、更に進んで學科成績相互の相關的研究を行ふとか、又は學科成績と他の測定結果との關係を研究しようとするが如き場合には、更に一層正確妥當で信賴し得る試驗を施行しなくてはならぬ。

このやうな要求に基いて、試驗問題の作り方、試驗の實施、採點の方法、評點の表現方法等の各方面に亘つて、從來の試驗方法を改善し、その試驗結果の評點たる數量的表示を一層合理的に處理し得るやうになつてゐる場合に特に之を學力測定と呼ぶ。測定の基點と單位とを明かにしてゐて、各兒童の學科成績をこれによつて測定し、その優劣を正確に表示し得るのが、學力測定の特徴である。測定の基點と測定の單位とを明かにしてゐる測定方法によつて兒童

の學業成績の優劣を表示することは、從來の試驗又は單なる檢査に於て缺けてゐるところである。從つて嚴密にいへば學力試驗と學力測定とは異るのである。こゝでは我々は學力試驗並びに學力測定の意味に就いて極めて一般的のことを述べたが、後に更に一層具體的に述べることゝする。

なほ國語科といへば通例は讀み方・綴り方・書き方の三分科を含むのであるが、こゝでは唯讀み方科を取扱ふこととし、而も小學校に於ける讀み方に限つて述べることとする。尤も時として他の教科教材に說き及ぶことがあるが、これは學力測定の意義を十分明かならしめようとするがためである。

二　從來の試驗法

兒童の學業成績評價の方法として從來から最もよく用ひられてゐるものは、口問口答法と論文試驗法である。今これ等の方法についてその性質と得失とを吟味してみよう。

口問口答法　學級の教師が口問口答法を用ひる主要目的は、兒童が如何によく理解してゐるかを見るにある。これと共にまた教師は恐らくは兒童をして熟考せしめ、興味を喚起し、發動的ならしめるために用ひるであらう。併しこの口問口答によつて兒童の學力を評價しようとするならば不完全な結果を得るに過ぎない。卽座的の口問による口答の評價は粗雜であり、不正確であり、從つて誤つた印象を與へ易い。同一の問題について全學級兒童に一樣に尋ね、一樣に口問させるといふことは不可能のことである。それに兒童の風采・容貌、親の社會的地位、平素の行動などによつて、その評價が左右され易い。なほ兒童によつては聲高くよくしやべるもあり、これに反して人前では內氣に遠

慮して口答する兒童もある。かゝる場合には答の適切さを決定する上に少からぬ影響を與へられるものである。

學力試驗としての口問口答は不正確で信賴し得ないものであることを教師はよく承知してゐるが故に、學力評價に際しては何れの教師も筆答試驗を用ひる。

論文的試驗　讀み方科に於て、漢字に讀み假名をつけさせるとか、書取試驗を行ふとか、語法上の誤りを直させるが如き試驗は、國史・地理・理科などの如き學科の試驗法とは形式上異つてゐたけれども、一般に知識學科に於ける試驗は論文的試驗であつたといへる。問題を若干出して生徒をしてこれに筆答させるもので、「比較せよ」、「論ぜよ」「説明せよ」、「理由を擧げよ」などといふ問題形式によるものが卽ち論文的試驗の代表的なものである。

一般に論文的試驗問題に於ては、多量の教材を自發的に整理組織して之を文章を以て表現しなければならぬ。長い間學級で讀んだり議論したりしたすべての事を、心的に回顧し、問題に關係せる事項を選擇し、そして問題の要點に答へるために、首尾一貫した表現をしなければならぬ。從つて論文的試驗は、兒童をして考察せしめ、自己の知識を整理組織せしめる價値を有するといふことは認めなければならぬ。併し乍ら他面に於てこの論文的試驗は幾多の缺點を有してゐることをも亦之を許さなければならぬ。卽ち、(一)答を書くのに多くの時間を費すので、提出し得る問題の數が僅少とならざるを得ない。廣く全體の教材に亘つて試驗することが困難である。(二)答案の採點に當つて正確公平な評點を與へることが頗る困難である。アメリカのスターク氏は中學校第一學年第三學期末に於ける一生徒の國語(英語)成績をそのまゝ石版刷にして、國語教師百四十二名に配布して採點させた。同一の答案を百四十二名の國語教師は如何に採點したであらうか。實にその採點の結果は百點滿點として五十點から九十五點に及んでゐて、七十點

以下の點をつけたものが十八人、九十點以上の點をつけたものが十四人あつた。この採點結果の蓋然錯差は四・八であつた。採點者が何れも同一點を與へれば蓋然錯差は零であるが、異る點をつけるがためにその値が高くなつて居るのである。異る教師が採點すればその評點が異るとしても、同一の教師は何時でも同一の點をつけると信じてゐる人が多い。併しこれさへも然うではない。教師が暫くの時間を經て答案を再び採點する時は、二度目の採點結果は多くの生徒に於て變つてゐるのである。論文的試驗の一大缺點は卽ち信賴し得る評點を付け得ない點にある。

論文的試驗は、教師の異る每に點數が著しく異り、同一の教師に於ても時によつて異るだけでなく、また少くとも最初には目的としてゐなかつた諸條件によつてその採點が影響され易い。文章の上手下手、誤字・當て字の多少、文字の美醜などによつて採點の影響を受けない教師は少いであらう。

それで論文的試驗に於て、その提出問題が最も適切であるならば、(一)材料を組織し思考する力を現はすことが出來る。(二)觀念・思想を明瞭に且有效に表現する技能を現はすことが出來る。(三)若し教師が注意深く採點するならば兒童の思考方法について價値ある洞察を得ることが多い。これ等はよい論文的試驗を課した際の長所であるが、併しこの種の利益を兒童に得させるためには、所謂試驗の際を利用しなくても、平生の授業に於て練習し得るところであり、殊に綴り方又は作文の練習に於てその效果を擧げ得るのである。

公平で信賴し得る評點を得る見地から見れば、最もよい論文的試驗であつても多くの缺點を有してゐる。卽ち(一)文字を書く速さ、文字の美醜、作文の上手下手の如き種々の無關係な要素が、生徒が教材を如何に理解してゐるかを評價しようとしてゐる教師に對して、影響を與へる。(二)上述のやうな缺點は避け得たとしても、客觀的採點基準が

確立して居ないから、或は寛大な點をつけたり、或は嚴格な點をつけたり、或は片寄つた點をつけたりして、公平適切な點をつけることが出來ない。(三)又、採點した點數だけではそれが、その年齢の兒童として果して如何なる位置を占めるのであるかを明かに知ることが出來ない。それは多數のものゝ成績から設定した尺度がないからである。その上に、(四)組織させ明瞭に表現させようとする問題を出すのであるから、勢ひ問題の數が少くなる。從つて試驗する範圍が制限され易い。

右に述べた第一及び第四の缺點を避けるために、新しく教師作製の客觀的檢査法が工夫されるに至り、第二及び第三の缺點を避けるために成績考査の規準が設定されるに至つた。考査規準のある檢査を標準化した檢査といひ、これによつて兒童の學力は始めて公平正確に測定し得るのである。

三　客觀的檢査法

論文的試驗の著しい缺點は客觀的に採點することが困難であり、書き方や作文の如き副次的の條件によつて採點が左右され易く、且問題數が少く試驗する範圍が比較的に限られた範圍となるといふやうな點にあつた。かゝる缺陷を除くために近來所謂客觀的檢査法が工夫されるに至つた。而もこの客觀的試驗法はこの外になほ若干の長所を有するので廣く用ひられるやうになつた。次にこの種の試驗問題の例を擧げてみよう。

〔一〕多數肢選擇試驗

各問題について上の方に示せる語に對して、下に示せる四つの中から最も正しい意味のものを一つだけ探し出し、その右側に

線を引きなさい。

喜捨——乞食。ほどこし。喜んですてる。差し出す。

しらべ——調子。たしかめる。しらべる。あはせる。

散果てる——木の葉がちる。秋の風が吹く。はてしがない。ちつてしまふ。

上例に於ては選擇肢は四つであるが、これは五つ又は六つにしてもよい。但し、三つ又はそれ以下は用ひないがよい。それは當推量によつて正答を得る率が多くなるからである。なほ、この多數肢選擇試驗の例として、早瀬松藏氏が小學讀本の全卷に亘つて、語の解釋・文の解釋・大意把捉・批評鑑賞・語法に類別して多數の問題を作られてゐる中の若干(卷六)を擧げてみよう。(「新應能教育の理論と實際」二〇〇—二一一頁)。

甲〔語の解釋〕

俵の山　(一)俵のやうな山、(二)俵の形につくつたこと、(三)俵をたくさんつんだ様、(四)俵をつんだ山、(五)俵をたくさんつんで山の形にしたもの。

ほう年　(一)今年のつぎの年、(二)新しく田ができたこと、(三)お米をたくさんとつた年、(四)作物の大層よくできた年、(五)去年。

新田　(一)お米のこと、(二)田の名前、(三)新しい田、(四)新しいお米、(五)大へんお米がとれた田。

稻こき　(一)稻をまくこと、(二)稻こぎといふものでこくこと、(三)うすに稻を入れてからをこきおとす、(四)稻のからをおとすこと、(五)稻のほをこくこと。

土間　(一)おもしのこと、(二)にはのこと、(三)いたのないところ、(四)家の床板のないところ、(五)いへのどまのにつてな

いところ。

もみをする　(一)うすの中に入れてこくこと、(二)もみをうすですること、(三)からをとる、(四)もみをかにかす、(五)もみをつぶすこと。

ときを作る　(一)鷄がコケツコーとなく、(二)にはとりがしらせること、(三)俵の山へ上つて鷄がなく、(四)時をしらせる。

乙(文の解釋)

うちでも土間に丸太をおいてその上に積んであります。

(一)うちでもゆかにきつたまゝのぼうをおいてその上につんであります。(二)うちでもゆかいたのないらうかに丸太をおいてつんであります。(三)うちでも土間に丸太をおいてその上に俵の山がつんであります。(四)うちでもゆかいたのない所にきつたまゝのぼうをおいてその上につんであります。(五)うちでもにはにつんであります。

俵の山へ上つてときを作りました。

(一)俵の上へ上つてこゑを上げました。(二)にはとりが俵の一番高い所へ上つてなくこと。(三)俵の上でなきました。(四)俵の山へ上つてこけこうこうとなきました。(五)俵の山へ上つてにはとりがこけこつこうとなきました。

朝晩めつきり寒くなりました。

(一)朝晩急に寒くなつた。(二)朝晩ほんとうに寒くなつた。(三)朝晩めつきり寒そうだ。(四)朝も晩もやたらに寒くなつた。(五)朝も晩もだん／＼寒くなつた。

丙(大意把捉)

私がたんぼへお湯を持つて行つてくると、おぢいさんが庭で腰をのばして「もうお晝かな」とおつしやいました。

(一)おぢいさんが腰をのばしてゐること。(二)お湯を持つていつたこと。(四)ゐなかのお晝のやうす。(四)たんぼへお湯をもつていつた。(五)私がたんぼへお湯を持つて行つてくると、おぢいさんが腰をのばして「もう晝かな」とおつしやいました。

臺灣の新高山さ。これは一萬三千尺からある、臺灣ではめつたに雪が降らないさうだが此の山のいたゞきはいつもつもつてゐるといふことだ。

(一)新高山の高さのこと。(二)雪のこと。(三)高い山はいつも雪がつもつてゐる。(四)臺灣のことで雪がいつもつもつてゐる。(五)新高山は高いから、いつも雪があること。

ふとんきてねたるすがたや東山。

(一)東山はふとんをきてねてゐるやうである。(二)人がねてゐるやうな山。(三)ふとんみたいな山。(四)東山のすがた。(五)ふとんきてねたるすがたや東山だ。

丁(批評鑑賞)

「今年はほんとうにほう年だ。今の分では去年より七八俵よけいに取れさうだ。」「さうです、新田が大へんよく出來ました。來年もやはりあの稲を作りませう。」

問　なぜ「來年もやはりあの稲を作りませう。」といつたのでせう。

(一)今こゝにないから。(二)今年はほう年だから。(三)今年大へんよく出來た稲だから。(四)たくさん米がとりたいから。(五)來年わるいといかぬから。

今日はうちの者がみんなたんぼへ稲こきに行きました。おるす居はおぢいさんと私だけです。

問　此の文をよんでうちの人々は皆どういふ人だと思ひますか。

(一)ひやくしやうをする人。(二)ゐなかの人。(三)よくはたらく人。(四)みんなしようじきな人。(五)たんぼのすきな人。

ふとん着てねたるすがたや東山。

問　どんな形をして居る山と思ひますか。

(一)小さい山。(二)けはしい山。(三)人がふとんを着てねてゐるやうな形。(四)のんびりした形の山。(五)すりばちのやうな形の山。

戊(語法)

去年より七八俵よけいに取れさうだ。

去年より七八俵よけいに取れた。

問　この二つのことばの○○のついてゐるところのちがひを書きなさい。

(一)取れるだらう、取れた。(二)取れるやうだ、取れた。(三)取れかけだ、取れてしまつた。(四)取れたかとれぬかわからぬ、取れたとわかつた。(五)人がいつたこと、自分がいつたこと。

「今年はほんとうにほう年だ、今の分では去年より七八俵よけいに取れさうだ。」「さうです、新田が大へんよく出來ました、來年もやはりあの稲を作りませう。」

問　「あの稲」といつたのはどの稲か。

(一)去年の稻。(二)來年つくる稻。(三)今年つくつた稻。(四)あの稻のこと。(五)よくとれた稻。

問 「ほんとうに」があるとないとでどうちがふか。

(一)去年まではとれなかつたが今年とれたことになる。(二)なければうそになる。(三)よけいとれたことになる。(四)いみがつよくなる。(五)ある方がよいから。

今どこのうちへ行つて見ても俵の山が出來てゐます、うちでも土間に丸太を置いて其の上に俵がつんであります。

問 此の文の中の「其の上」とは何の上か。

(一)土間の上。(二)丸太の上。(三)俵の上。(四)うちの上。(五)俵の山の上。

なほ語の意味檢査としては、

書物は、(一)著るもの (二)讀むもの (三)食べるもの (四)航行するものである

といふやうな形式で提出して、この文を正しいものとするに必要な「讀むもの」の右側に線を引かせることも出來る。

〔甲〕

商

人
泣
標
車
止
用
業

〔乙〕

書物を　拜觀する
危險を　集錄する
寶物を　豫防する
食物を　整理する
言行を　選擧する
　　　　栽培する
　　　　養成する

〔二〕 組合せ試驗又は結合試驗

「商」と結びつけて用ひられる字を、商の下にある字の中から探し出して、その間に線を例(商—人)にならつて引きなさい(上揭〔甲〕參照)。

上にある各の言葉と結びつけて正しい意味を表はすことの出來る言葉を下から選び出して、その間に線を例(書物を—整理する)にならつて引きなさい(上揭〔乙〕參照)。

組合せに用ひる項目數は、普通十乃至十五以下であつて、それ以上は用ひない。

〔三〕完成試驗

尋常小學國語讀本卷三

ハルノ日ハ〇〇〇〇デス。

センセイト〇〇〇〇ニイキマセウ。

モウ八十二日バカリ〇〇トナツヤスミデス。

私ハセンセイトベンタウヲタベルノガ〇〇〇〇デス。

中になにがあるか〇〇〇ごらんなさい。

天人のはごろもなら〇〇〇〇かへすことは出來ません。

これは正しい意味の文となるやうに、〇の中に假名を入れさせる檢査であるが、餘程注意して用ひないと、いろいろの答が出て來て採點に迷ふことがある。

〔四〕指圖試驗

柄のない林檎の繪を描いておいて、「コノリンゴニ　エヲ　ツケナサイ」の文章を讀ませて林檎に柄を書き入れさすが如き、毬の繪を描いた側に、「このまりに十の字のしるしをかけよ」と書いておいて、その指圖通りにさせるが如きは一例である。主として低學年の讀み方考査に於て用ひられる。

從來から行はれてゐる假名文字を漢字に直す書取檢査、漢字に讀み假名をつける讀みの試驗の如きが、こゝに述べ

た客觀的檢査の一種と見てよいことは云ふまでもない。

これで讀み方科の試驗に用ひられる問題例の擧示をやめるが、これ等は完成試驗を除く他は何れもその採點が客觀的である。そして文字を書く必要がない。完成試驗の際には文字を書くけれども極めて簡單である。そこで書字の速さや性質、作文力の如き要素の關係するところが殆んど無い。兒童は短時間にかゝる問題の多數に答へ得る。從つて論文的試驗よりも一層廣い範圍の材料を以て試驗することが出來る。

客觀的檢査にはこの他になほ長所がある。客觀的檢査は論文的試驗よりも採點が容易で著しく客觀的であるから、敎師は異る學級の成績を比較し得、校長は全校を通じて同じ檢査を與へることが出來る。またかゝる檢査は敎授上から見ても一つの長所をもつて居る。檢査成績を採點してから、敎師がその答案を各兒童に返してやる時は、各兒童は自分の誤を調べてみることが出來、疑問があればそれに就いて敎師に尋ねることが出來る。論文的試驗であると、低い點を與へられた兒童が、自分の答案について誤つてゐるところを明かにし得ないでおくのが普通である。然るに客觀的檢査では誤つてゐるところを正確に兒童に知らせることが出來る。かういふことは從來殆んど顧慮されなかつたことではあるが、敎授上重要なものといはねばならぬ。

いま敎師作製の客觀的試驗の長所を列擧してみれば次のやうである。

(一)採點が客觀的である。

(二)書字・作文の上手下手の如き副次的の條件が除去される。

(三)多數の問題を提出することが出來、而も試驗に要する時間は極めて短くてよい。

（四）採點が容易である。
（五）客觀的に採點し得るから、學級・學年・學校相互間の比較が出來る。
（六）採點が客觀的であるために、兒童をしてその答の誤つてゐる點を明かならしめ得る。

このやうに客觀的檢査は價値があるので、漸次廣く用ひられるやうになつたが、更に問題の選定に意を用ひ、提示の方法を改善するならば、一層有效にして而も便利になり、益〻廣く利用せられる樣になるであらう。

四　單なる試驗の他の缺陷

口問口答試驗・論文的試驗・客觀的檢査などが、現在多くの學校に於て學力試驗として廣く用ひられてゐるといつてよい。これ等の試驗は兒童の學力を評價し、その低劣原因を探求する上に、それぞれ長所と短所とを有してゐることを述べたが、併し、これ等を眞に學力試驗として學級敎師が用ひる際には、なほ共通的に幾多の缺陷を示すのである。私共はこゝにその共通的缺點について述べてみよう。

（一）敎師の與へる評點の分配狀態　評點の分配狀態とは同一組の兒童の一定學科に於ける成績について一々の點數を得た人員を調べて、全體としてどんな狀況に分布して居るかを見ることである。同じ組の兒童といつても入學に當つて、智能について嚴密な規準を設けて選擇した場合とさうでない場合とでは點數の分配狀態は異る。そこで先づ智能上何等の規準も設けない場合、卽ち無選擇兒童の場合について考へる。無選擇兒童の場合をよく代表するのは普通の町村小學校に於ける如きもので、多くの師範學校附屬小學校の如き場合は餘程選擇されて居ると考へられる。

さて多數の無選擇兒童の一定の性能の現はれた成績、例へば算術・國語・圖畫等の成績について考へるに、それ等の成績の優劣は、一方には各兒童の身體の狀況・智能・興味・努力等による差異と、他方には教育法、家庭の教育に對する關心の度等の廣い意味の環境との合成的結果である。その中でも智能の如き個人的特質が特に重要である。そして智能は主として遺傳的のもので、これに依存することの多い學業成績の分配は、從つて物が偶然の機會によつて現はれる場合と同じ様になると思はれる。偶然の機會に支配されて起る現象の分配狀態は、蓋然分配となる。これを圖示したものを蓋然曲線といふ。この曲線の特徴は中央に於ける頻數が最大で、それから左右に進むにつれて頻數を減ずること、而してその減じ方は初め徐徐に、次に急激に、そして再び緩徐になること、而して中央から左右に進むとき、その減じ方が左右共に等しく、所謂對稱的であることとである。

一定の學科成績もその檢査法が信頼し得るものであり、被驗者が無選擇のものが多數である時には蓋然曲線的の分配をなすであらうと思はれる。然るに多くの學校について、その評點或は評語の分配狀態を調査してみると、この正常的分配から遠ざかつて居ることが多い。シカゴ大學の附屬中學校で調査したものによると、甲・乙二人の教師が調査した國語の評語の分配は次の表の通りである。表中、Aは最優、Bはそれに次ぎ、C・Dは更にこれに次ぎ、Eは落第を示す。

評語	A	B	C	D	E
甲教師	7.5%	22.0	47.5	16.0	8.0
乙教師	36.0%	30.5	24.0	6.0	4.5

右の表によると、甲教師の與へた評語の分配は略々蓋然曲線的であるが、乙教師のは劣等の方に頻數が稍少で、成績優等な方に向つて順次頻數を増し、Aの評語を得たものが最も多いことになつて居る。右の二人の教師の評語の結果の中で甲教師のが比較的正確であつたらうと思はれる。但し中學校では多少選抜された生徒であるから、その成績は正常な蓋然曲線を描くとは思はれないが、それにしても、乙教師の場合の如きものは實際を現はさないものであらう。普通の小學校などで、各の評語又は評點に應ずる頻數の割合は、大凡、次の如き標準によつて大體の見當をつけて、調査の結果を修正すればよいと考へられる。

評語	甲	乙	丙	丁	戊
頻數	7%	24%	38%	24%	7%

これは五段階に分けた場合であるが、若し三段階に分ける場合には、甲が一六%、乙が六八%、丙が一六%となるやうにすればよい。

〔二〕考査基準の高低　教師によつて採點が甘かつたり、辛かつたりしては考査の正確を期することは出來ない。甘い點をつける教師の與へる點數の平均は大きく、辛い點をつける教師の與へる點數の平均は小さい筈である。點が甘いといふのは考査の基準が低いことであり、辛いといふのは基準の高いことである。

米國のグレーは數學の答案十一枚をとつて六名の教師をして採點せしめ、各教師每にその與へた點數を平均して次の如き結果を得た。

教師	A	B	C	D	E	F
平均點	78.7	74.0	61.4	65.5	75.5	58.0

右の表によればA教師の與へた點が最も甘く、F教師のが最も辛い。そしてその最高と最低の差は實に29.7點である。數學の如き學科の採點に於て、同一の答案を採點するに際し、教師が異るためにこれ程の差異を生ずることは最も注意すべきことといはなければならぬ。

グレーは又六名の教師をして國語の答案二十枚に採點せしめ、各教師毎に平均を出して、次の表に示すが如き結果を得た。

教師	A	B	C	D	E	F
平均點	80.3	83.7	78.5	54.0	70.0	79.5

こゝではB教師が最も甘い點をつけて居り、D教師が最も辛い點をつけてゐて、その差は實に29.7點、即ち約三十點である。右二つの例で教師の考査基準には如何に著しい高低の差があるかを認めることが出來る。若しも一定の教師は何時でも考査規準が高いとか、低いとかいふことが一定して居れば、その高過ぎたり低過ぎる程度を測定して、考査した結果に割増又は割引をして多くの教師間の基準を一致せしめることが出來る。然るに實際に於ては、同一教師が考査するに當つて、或る答案に對しては甘く、他の答案に對しては辛いといふやうなことがあつて、單純に割引又は割増をすることは出來ないのである。その一例としてスタークの研究を記すこととする。

スタークは地理の答案三枚をとつて、九十一名の教師をして、滿點を二十點として採點せしめて、總平均點を計算したのに、甲答案では16.1點、乙答案では10.9點、丙答案では0.1點であつた。然るにその答案をとつた學級の受持教師は右三つの答案に對して、順に12.0點、16.0點、2.0點を與へて居る。若しも九十一名の教師の與へた平均點は略

々その答案の眞の値を示して居るとすれば、右の受持教師は甲答案の場合には、辛い點をつけたことになり、乙答案及び丙答案には甘い點をつけたことになつて居る。このやうに考査規準が動搖してはその採點結果の客觀性を失はしめるといはなければならぬ。

〔三〕同一成績物に對して與へる點數の不一致度　同一の成績物に對して、多くの教師が評點を與へた場合に、その結果は一致すべき筈であるのに實際に於ては一致しないのが普通である。その不一致度を一にまた脫逸度といひ、その指數として平均錯差・四分錯差・標準錯差を用ひる。こゝでは平均錯差の計算法だけを述べておかう。平均錯差といふのは全體の平均と個々の値との差の平均である。例へば今一つの答案に對して十人の教師が評點を與へて、次の表に評點を記してある行に示すが如き結果になつたとする。即ちA教師は75點、B教師は90點、かくてJ教師の90點までそれ〲の點數を示してある。これ等十人の教師の與へた點數を平均すると平均の欄に示す通りに82點となる。この平均82點と各教師の與へた評點との差を求めると、A教師の場合は −7、B教師の場合は +8、C教師の場合は −2 といふやうになる。一々の場合は表中「平均との差」の行に示す通りであつて、或る教師は平均よりも大きい點をつけ、他の教師は平均よりも小さい點をつけてゐる。即ち平均から脫逸して居る。そしてその脫逸の度には大小がある。そこでこの一々の脫逸の程度を一つの代表的な値で現はさうとするには、これ等の一々の脫逸の平均を求めればよい。それにはプラス・マイナスの符號を除いて、絕對數についての平均を求める。これを計算すれば右の表の平均の欄にあるやうに5.6を得る。

教師	評點	平均との差
A	75	−7
B	90	+8
C	80	−2
D	85	+3
E	88	+6
F	70	−12
G	85	+3
H	77	−5
I	80	−2
J	90	+8
平均	82	5.6

これが平均錯差である。

平均錯差を求めることによつて、同一の成績物に對して教師の與へる評點間には割合に大きい差があることが、今までに多くの研究者によつて示されて居る。

前に述べたグレーが、十一枚の數學の答案に、六名の教師をして採點せしめた結果について、各答案毎に平均錯差を求めてみると次の表のやうであつた。

答案	平均錯差
1	8.0
2	4.2
3	3.1
4	9.2
5	14.7
6	3.0
7	7.2
8	7.3
9	10.2
10	4.2
11	7.0
平均	7.1

これによつて見ると答案によつて、平均錯差には大小があつて、或る答案では各教師の評點が割合に一致し（平均錯差の小さい場合、即ち答案6・3・2・10の如き）、他の答案ではそれが割合に一致して居ない（平均錯差の大きい場合、即ち答案5・9・4の如き）。このやうに答案の種類によつて、評點の不一致を異にするが、前表に於ける十一個の平均錯差の平均を求めると7.1となる。

同じくグレーが二十枚の國語の答案に、六名の教師をして採點せしめた結果は次の表に示すやうであつた。

答案	平均錯差
1	9.6
2	9.3
3	7.5
4	10.8
5	6.7
6	11.8
7	10.8
8	7.6
9	7.6
10	5.3
11	12.0
12	6.0
13	11.2
14	8.2
15	10.5
16	8.3
17	9.5
18	16.8
19	8.5
20	5.4
平均	9.2

この表による時は答案10が最も評價し易くて、答案18が最も困難であつたと見える。二十個の平均錯差の總平均は9.2で、前の數學に比べれば國語の方が評價するに困難であつたことを示して居る。

右と同じやうな結果は文學士福富一郎氏の研究に於ても亦之を見ることが出來る。氏は讀方・算術・國史について大連及び奉天の小學校第六學年兒童に特別の試驗を試み、その答案中から讀方・算術に於て各六枚、國史に於て三枚

の答案を選び、これ等を同一人の手によつて一定の形式の下に謄寫に附し、五十四名の教師に採點上の注意書を與へて嚴密に採點せしめた。同一人の手によつて謄寫せしめたのは筆蹟の美醜、紙面に於ける文字の位置、その他成績としては副次的な總べての要素を出來るだけ除去するが爲であつた。

國語			算術			國史		
答案	平均點	平均錯差	答案	平均點	平均錯差	答案	平均點	平均錯差
1	88.72	6.02	1	96.61	3.47	1	92.51	6.77
2	87.77	6.71	2	87.35	12.02	2	78.57	16.67
3	84.96	7.67	3	65.48	22.94	3	44.42	8.42
4	76.92	9.57	4	40.62	27.98			
5	68.18	6.31	5	6.01	8.54			
6	63.73	8.20	6	5.57	2.91			
平均		8.08	平均		12.91	平均		10.62

五十四名の教師が採點した結果について各答案毎に平均錯差を示せば上表のやうである。この表には各答案の得た平均點をも記入してある。こゝにも答案の性質によつて脫逸度に大小があることが示されて居る。而して今までの研究を綜合してみるに成績の極めて優れた答案と、極めて劣つた答案では割合に脫逸が小さく、之に反して中位の成績物に對しては評價する人々の意見の一致することが少く、從つて平均錯差によつて示される脫逸度は大になる傾向が著しい。そして普通の學校に於ける成績物の大多數は中位の成績のものであるを思ふとき、各教師の與へる評價の不一致は成績考査上に於て重大な問題である。

(四) 多數の答案について優劣の差をつける精密さ　前に述べた脫逸度の大小によつて、評點の與へ方の正確さを見ることは一面からいへば前に述べた評價の基準の高低を見ることと關係して居る。即ち或教師は甘い點をつけ、他の教師は辛い點をつけると、その結果は脫逸度を大ならしめる結果になるのである。然るに多數の成績物に評點を與へるときその平均點から見ては甲の教師も乙の教師も略々等しい點數を與へて

居り、從つて見かけでは評價基準は等しいとしても、それだけではまだ評價が正確に出來て居るとは斷定し得ない。そこで(三)に述べた脫逸度を調べる必要があるのであるが、併し脫逸度を調べてその値が小さいといふことが分つても、なほ他の方面から來る評價の不正確さについては之を知ることが出來ない。それは與へた點數の分配範圍が極めて小さいならば、成績について正確な評價をしたことにならないからである。詳言すれば、普通の學校に於ける兒童生徒の成績考査の場合のやうに、實際に著しい優劣の差のある多數の成績物に對して與へた點數が、七〇點から九〇點位の範圍に及んで居るだけであつて、精密に優劣の差がつけられて居ないならば、たとひ多數の敎師に採點させた評點の脫逸度は小さくても、その考査結果は正確妥當であるとはいはれない。蓋し正確な考査といふのは優れたものを優れたとし、劣つたものを劣つたものとして評價し、而かもその優劣の差が明かに示されたものでなければならないからである。その點から見て、從來の成績考査の結果は甚だ不滿足なもののやうである。それについては各學校で各學科每に一人々々の敎師の與へた點數或は評語の分配の範圍を調査すれば明かになるであらう。それは(一)に於いて述べた評點の分配狀態を調査すると同時に調べられる事柄である。

田中は嘗て東京高師附屬中學校に於ける入學志願者選拔考査の際に書方成績をとり、これを五名の審査員をして一點から一〇點までの評點をつけさせたことがある。受驗者の書方成績數は四四二枚であつた。各審査員たる敎師の與へた評點の分配は次の表の通りである。こゝで中間數といふのは小さい點の方から數へても大きい點の方から數へても總數の丁度中央に當る成績物の得る點數である。それ故に大體に於て中間數は平均と考へておけばよいのである。この表を通覽するに、何れの敎師の與へた評點も皆最小の一點から最大の一〇點にまで及んでゐて、割合によく優劣

評點	教師				
	甲	乙	丙	丁	戊
1	35	45	61	12	146
2	39	163	103	33	61
3	61	69	32	89	25
4	42	45	75	62	66
5	51	58	52	.35	51
6	73	33	46	61	16
7	38	8	24	94	47
8	44	11	24	21	16
9	49	7	19	22	9
10	10	3	6	13	5
中間數	5.86	3.19	4.33	5 71	3.56

の區別がつけられてゐるといへる。併し中には六點以上の頻數が極めて少いものがある。教師乙と戊の二人の場合はその著しいものである。從つてこの二人の中間數は他の教師のに比して著しく小さいものになつて居る。即ちこの二人の教師の採點は優れた成績物についての區別が割合に不十分であつたといへるであらう。

〔五〕**異る教師の與へた評點間の相關關係**　異る教師が同一成績物を採點してもその評點がかなり異つてゐることを述べたが、それは脫逸度による調査であつた。この一致不一致度はまた評點間の相關關係を調べることによつても明かにされる。相關關係とは之を簡單に定義すれば、二つの現象間に於ける相應又は一致の關係をいふ。例へば溫度が高まれば水銀柱の長さは增加し、反對に溫度が下れば水銀柱の長さは減少する。即ち一方の現象が增減すれば他の一方の現象も亦それにつれて增減する場合の如きは相關關係のある一つの場合である。又光の强さは光源からの距離の自乘に反比例して增減する。即ち光源からの距離が大になれば光度は反對に減少し、之に反して光源からの距離が減少すれば反對に光度は增加する。このやうな場合も亦二つの現象間に相應の關係即ち相關關係があるといふ。そして前者の例の如きは極積的相關關係があるといひ、後者の例の如き場合を消極的相關關係があるといふ。若しも右二種の如き關係がないときには兩者の間に何等相關關係がないといふ。この相關關係の程度は一定の公式によつて求めることが出來る。その公式にピアスンの**錯差積法**、スピアマンの**列位差法**、**相關比**などの種類があるが、何れにしても積極的相關關係が完全であればその

値は＋1.0となり、相關關係が無ければ０となり、消極的相關關係が完全であれば－1.0となる。

一定の公式を用ひて相關關係の値を求めたならば、その値の信頼度を計算してみなければならぬ。即ちその相關値の蓋然錯差を他の公式によつて算出しなければならぬ。そして相關値が、その蓋然錯差の三倍又は四倍よりも大である時には、その相關値は信頼し得るのである。

前述の東京高師附屬中學校の入學考査のときに得た書方成績四四二枚に對して五人の教師が與へた點數間の相關關係を求めてみれば次のやうである。この表中で相關値の最小は教師甲と丙との間に於ける0.426でその蓋然錯差は0.026であるからこの相關値は十分に信頼し得る。從つて他の相關値も亦全部信頼し得る。一般にこの種の相關値としては餘程高い値であるけれども、それは比較的のことであつて、同一成績物に對してかくもまち〳〵の評點が與へられるのかともいひ得るのである。理想的にいへば〇・九以上の相關値を得るやうに採點法を改善しなければならないのである。

五人の教師が442枚書方成績を考査して採點した點數間の相關値

教師	甲	乙	丙	丁	戊	平均
甲	——	.471	.426	.572	.522	.498
乙	.471	——	.526	.578	.547	.531
丙	.426	.526	——	.479	.487	.480
丁	.572	.578	.479	——	.610	.560
戊	.522	.547	.487	.610	——	.524
平均	.498	.531	.480	.560	.542	.522

〔六〕**配當點數の不公平**　學校に於ける試驗問題は一時に數問又は數十問提出するのが普通である。この際は各問題に配當する點數が公平でなければならぬ。即ち一層むづかしい問題には大きい點を與へ、比較的やさしい問題には小さい點を與へるのが公平である。然るに考査の實際に當つてはこの公平の原理が無視せられて居ることが多い。試驗問題が五つであれば、その難易に關係なく各問題に對して總べて二〇點宛と

することもあれば、むづかしさを考慮して配當點に差等をつけるとしても、それが極めて主觀的である。例へばこゝに、(1) $3+5$ と (2) $17+8$ の如き二個の加算問題があるとする。誰でも第一問よりも第二問の方が一層困難であることは分るが、然らば(2)は(1)の何倍むづかしいかは多くの兒童について實驗してみなければ正確には答へられない筈である。それを勝手に同價値に見るのも又勝手に(1)を一點とし、(2)を二點といふやうにきめることも共に間違ひである。各問に配當する點數が異れば、たとひ各問についての考査は極めて正確に出來ても、總合の結果は極めて異つたものになる。

これに就いて田中はかつて一つの調査をしたことがある。それは某中學校の入學試驗の算術問題に對する二一八名の答案について、一方ではその學校で適當と考へる配當點によつて採點した結果による順位を定め、他方では各問の困難度に應ずる正當な配當點を算出してこれによつて採點して各受驗者の成績順位を定め、この二つの順位の差を求めたのである。その差の零のものは二一八名中僅かに一九名で、差が一のものが二〇名、二番の差を有つたものが二二人といふやうであつて、最後に三十四番の差を生じたものが一名まであつたのである。このやうに各問題に配當する點數の差異によつて、成績の順位は著しく狂ひを生ずるものであることを知らなければならぬ。困難度に應じて配點するには、先づ各問毎に正解者の百分比を求め、これを一定の表によつて標準錯差量に換算し、その標準錯差 -5 を0點とし、-4 を一點とし、-3 を二點とし、順次かくして $+5$ を一〇點として用ひるのが最もよい方法である。

（田中「教育的測定學」四〇六頁以下參照）

〔七〕性能の分析の不充分　從來の成績考査では一定學科に於ける性能を分析的に考査することが不十分である。卽

ち讀方成績といつても、その中には種々の要素的性能がある。例へば早瀬松藏氏が讀方考査の方面として擧げたところによれば、(一)文字の讀み、(二)文字の書取、(三)熟語構成、(四)類似語及び反對語、(五)語の應用、(六)文章の解釋、(七)文章の大意把捉、(八)文章の批評鑑賞、(九)語法等がある(早瀬「讀方考査に於ける評定のつけ方」教育心理研究第四卷五三〇頁——五三八頁參照)。兒童の讀方成績の劣等である原因探求の場合には、これ等のものは別々に考査しなければ、低劣原因の診斷に役立たない。次に讀方の一般的成績を見る際には、これ等の要素に對して、それ〴〵適當の重みを與へて、その重みに應じて採點し、また合計點を求めなくてはならぬ。

〔八〕速度の無視　右に述べた性能の分析は主として性質的方面からいふのであるが、次に考査に當つて考へなければならぬ點は速度の方面である。例へば一つの問題を成し遂げるのに十分間を要したものと十五分間を要したものとでは、たとひ答案の出來榮えは同等であつても、その力は著しく異つて居ると見なければならぬ。然るに從來はこの速度の方面が無視されて居るのが普通であつた。即ち與へられた時間内に、やつと出來たものも、餘裕をもつて答案を書いたものも共に一樣に取扱つたのである。

〔九〕客觀的基準の缺乏　單なる成績考査の一大缺點は、考査の結果を解釋して兒童生徒のそれ〴〵の學科に對する成績の程度を明かにする爲の客觀的標準卽ち代表値の無いことである。今までの普通の考査では、精々その組の中に於て各兒童が占める位置を明かにし得るだけで、それが例へば他校の兒童に比較して如何なる地位にあるか、全國的の尋常六學年兒童として如何なる成績を示すかは之を明かにすることは出來ない。それを明かにするには廣い範圍に亙つて同一問題で考査した全國的の兒童の成績を代表する如き代表値又は規準成績を設定して置かなければならぬ。

かゝる代表値のある檢査を標準化した檢査といふ。このやうな檢査によらなくては各兒童の成績の優劣を正確に測定することが出來ない。從つて學力測定運動はこの點を一つの重要な目標として居るといふことが出來る。

五　標準化した檢査

上述のやうな諸種の缺陷を出來るだけ避け、各兒童の成績を解釋するための規準を設定した試驗が即ち標準化した檢査である。これは代表値を有する檢査といつてもよく、これによつて始めて我々は各兒童の成績を正確に測定し得て、成績優劣の原因を明かにすることが出來、從つて成績低劣に對する救治策を講じ得るのである。

次に代表値のある學力檢査の特色を述べることとする。尤もこれから述べる學力測定の特色は現在世の中に在る標準化檢査の總べてのものが有してゐるといふのではなくて、むしろかういふ特徵を有つやうに努力してゐるといふべきものである。

〔一〕客觀的規準のあること　各兒童の學力の優劣を評價する際には必ず一定の規準がなければならぬ。從來の試驗ではこれが全く各敎師の主觀的規準によつた。經驗の多い敎師の場合ではその主觀も餘程客觀的のものと一致するであらうが、それでも別々の試驗を行つてゐては學級成績相互の比較、又は學校相互の比較の際にその評價が困難である。敎師の主觀による時は、或は採點が寛大であつたり或は嚴格であつたりして正確な評價を得ることは期し難い。このやうな多くの弊を避け客觀的に評價するには一定の規準を設定しなければならぬ。規準はまた代表値ともいふ。代表値は性別・學年別又は年齡別に示された平均點と標準錯差とである。我々はこの代表値によつて同一の試驗を受

けた個々の兒童の學力を測るのである。この代表値が設定されてゐるといふことが學力測定の一大特色である。代表値は、兒童の智能、その家庭狀況などを考慮して、最も代表的のものを數百名選び、これに一定の試驗問題を課し、合理的に採點して、始めて得られるものである。

〔二〕試驗問題の作製が一層合理的であること その爲に少くとも一致性・信賴性・分類性の三原理を考慮して問題を作製する。

〔1〕**一致性** こゝに一致性といふのは測定する性能が常に同一であることとである。この目的の爲に測定の問題を作る時には、最初にその學科の目的又は要旨を心に浮べてその要旨に添ふやうな問題を選ぶ。次には何を査定するかを定立する。文字の書取成績を調べるとか、又は文章の大意把捉の成績を考査するとかといふやうに一定の限界を定めねばならぬ。一度考査する方面が定まつたならば、次にはその査定する性質だけを測定する試驗問題を作る。これを或は本質的の問題を選ぶといつてもよい。測定の目的が敎授せる知識及びそれについて考へる能力の測定である場合には、人物の道德的特質又は文章の形式等の如き直接に關係するところなき事項を測定するやうなことがあつてはならぬ。若し我々の試驗が從來行はれて來たやうな主觀的のもので、算術の試驗に於て、或る問題はかなり困難な讀み方の試驗となつて居り、國史の試驗が作文の試驗と殆んど異るところがないならば、それは一致性に乏しいものといつてよい。

本質的の問題を選ぶといつても、唯一人の敎師が問題を選ぶ時は、ともすればその內容が一方面に偏し易い。偏した問題で試驗しては兒童の力を十分に見ることが出來ない。この缺陷を補ふ爲には二人以上の同一學科の受持敎師が

別々に選んだ問題を合せ用ひるやうにするがよい。

從來行はれた學力試驗は、殆んど大部分が力量試驗であつたといへる。併し作業速度も重要な一方面であるからこれを考慮することが必要である。いろは歌を普通の順に唱へることは迅速であるが、これを逆にいふには隨分時間を要する。正確にいふことはいへるけれども非常に時間がかゝる。これは練習を積んでゐないからである。この樣なことが、讀み方に於ても了解に於てもある。作業速度の遲いものには更に反復練習を課する必要がある。

〔2〕**信頼性**　信頼性の無い試驗は、試驗としての價値が無いにも拘らず、從來はこれについてあまりに無關心であつた。信頼性は普通に自己相關關係といはれ、試驗それ自身との一致を意味する。同等の試驗を同一の兒童に二度施して、各兒童が兩回とも同じ點數をとるならばその試驗は全く信頼し得る。信頼性を調べるには(一)一度試驗を施行して、それからその記憶的影響が除去される程の間隔をおいて、再び同一の試驗を課し、そして二回の試驗の結果によつて相關關係を求める。(二)同一の試驗を例へば奇數番號の問題と偶數番號の問題とに二分して、その相互の相關關係を求める。(三)同種の學力檢査系列が二種ある時は、これを同じ生徒に課して得た二つの點數からして相關關係を求める。かうして得られた相關値を信頼係數といふ。信頼係數は少くとも〇・八以上なくてはならぬ。

信頼性を増すにはどうすればよいか。この爲に第一には客觀的に採點し得る問題を作るのである。前に述べた客觀的檢査の發達するに至つたのはこれが爲である。第二には一試驗問題の數を多數にする。信頼係數〇・九以上を得る爲には、內容上代表的に選ばれた問題が六十乃至八十以上であることを要する。問題數が多ければ試驗時間が長くかゝるやうに思はれるが、客觀的檢査を用ひれば極めて短時間に多數の問題に答へ得るのである。

〔3〕 **分類性**　信賴性を增す爲には、更に分類性について考究せねばならぬ。問題數を增しても、分類性について考へるところがなければ、これを增さない場合と同じ結果になるからである。極端な例をとつていへば受驗者の何人も解し得ない問題、又は反對に全部のものが正解し得る問題は、それが何程多く提出されても事實信賴性に關係しないことは明かである。測定される學力に量的差異がある兒童群を檢査して、その點數に差異がなかつたならば、その測定用具たる試驗問題系列は分類性を缺くものといはなければならぬ。或る試驗が果してこの分類性を有するか否かを調べるには次のやうな規準による。從來の研究の結果によれば、(一)普通の學校の相繼げる學年の兒童で特別に選拔されたものでない場合には、一般に學年の進むにつれて平均點の增進を示すものである。(二)若し數百人又は數千人の兒童について妥當な試驗を施行するならば、その點數の分配は所謂蓋然分配に近似して居る。この二つの假定を標準として分類性を調べるのを普通とする。

普通の兒童群である時は、その學力に優劣の差があるのであるから、これに課する力量試驗の問題は、各問題の困難度を、平易なものから順次困難なものへと、殆んど等量に漸進させておくことが肝要である。

〔三〕**檢査實施手續の一定して居ること**　兒童が一つの試驗に答へて出來た成績は、多くの條件に依繫して居る。例へば、文字又は數字を書く速度、生徒が成績をよくしようとする努力、生徒の身體的情態、日時の如何、試驗官の態度、生徒に與へる說明及び指示、試驗の形式、その他の些細な條件が皆兒童生徒の學業成績に影響を及ぼすものである。このやうに兒童の試驗成績は數多の條件によつて影響を受けるものであるから、我々はこれ等の條件中、統御し得るものは、常に一定的ならしめるやうに努めなくては、信賴性の高い結果を得ることが出來ない。例へば口頭試問

よりも試驗問題を板書する方がよく、板書よりも印刷に附して問題を各兒童に與へる方が一層よいといふが如きは、その條件を單一ならしめるが爲である。

受驗中に鉛筆が折れて、十分に實力を發揮し得ないことも往々ある。それ故に鉛筆は二本以上用意せしめておくがよい。また新しい試驗形式の問題である時は、その解答形式によく熟知してゐるものと然らざるものとがある。かゝる際にはその方法を丁寧に說明し、更に若干の練習問題を提出する必要がある。學力測定がこれ等の點に十分に注意してゐるのは從來の試驗よりも異るところである。

〔四〕**採點方法の合理的であること** 試驗成績の點數は相互に比較し得るものでなければならぬ。算術の七點は地理の七點と同樣な意味を有するものであり、また理科の七點と同樣の意味のものでなければならぬ。また八點と九點との差の一點は、五點と六點との差の一點に等しいものでなければならぬ。その爲に我々は各學科の採點にあたつて等しい測定の基點と測定の單位とを用ひねばならぬ。

如何なる測定でも、それが科學的に取扱はれる場合に、最初に必要なことは、その測定で用ひられてゐる單位の意味を明かにし、且その計算の始められた零點卽ち基點を明かにすることである。現在我々が用ひてゐる華氏寒暖計の歷史及び華氏寒暖計が攝氏寒暖計に對する關係は、零點を定位することの如何に重要であり、單位の性質を知ることの如何に重要であるかを明かにするであらう。嘗ては溫度に於ける絕對的零點卽ち總べての分子の活動の止む點をば純粹の水が凍る點以下華氏三十二度と思つた。今では溫度上の絕對的の零點は攝氏氷點以下二百七十三度と思はれてゐる。そこで現在用ひてゐる華氏及び攝氏の零點は絕對的のものでなくて、人爲的零點である。教育的測定に於て絕

對的零點を精密に知ることは困難である。それで普通には正常の人々が容易になるほどと思へるやうな零點（基點）を決定すればよい。測定單位についても同様である。かくて今日普通に用ひられるに至つた測定の基點は平均點であり、測定の單位は標準錯差である。

平均點を測定の基點とし、標準錯差を測定の單位として用ひるのが最もよいのであるが、この場合プラス・マイナスの符號を用ひなければならぬ。そこで我々はマイナス五標準錯差を零點とし、平均を五〇とし、プラス五標準錯差を一〇〇とするやうな點數を用ひることとし、之を偏差値と名づけてゐる。偏差値を用ひればプラス、マイナスの符號を用ひる面倒がなく、また小數點を用ひる場合が少くてすむ便益がある。

偏差値は答案を一度採點した後にこれを換算して算出するものである。それで採點して得たまゝの點數を粗點數といふ。代表値としての平均並びに標準錯差はこの粗點數から計算して得たものである。この平均及び標準錯差を基礎として、次の公式を用ひる時は、各兒童の粗點數は偏差値に換算される。

$$\text{偏差値}=\frac{\text{各自の得點}-\text{平均點}}{\text{標準錯差}\times\frac{1}{10}}+50$$

偏差値を用ひる時は、點數を相互に比較し得、また正當に合計點を算出し得る。

一般に測定單位の精粗は、その試驗の信頼性に關係する。少數の被驗者から得た結果に對して、精密な單位を用ひることは無意義であるが、これと共に、またあまりに粗大な測定單位を用ひる時は、生徒本來の分類性を妨げることになる。即ち用ひる單位が比較的に大きい時は、多數の生徒がたとひ學力の差異があつても同點をとることになる。

學力測定に於て偏差値を用ひる時は、多數の受驗者を凡そ五十段階に段階づけることが出來るのである。

合計點を算出する際には、合計點に及ぼす各要素の重みが問題となる。合計點を左右する條件は何であるかといふに、多くの人々は絶對的點數卽ち粗點數の大小によると思つてゐるやうであるが、事實は決して然うでなく點數の擴がりの度によるのである。卽ち脫逸度の大小によつて合計點は左右される。脫逸度は平均錯差・四分錯差・標準錯差を以て示すが故に、合計點を求める際には、各要素毎に四分錯差又は標準錯差を求めなくてはならぬ。標準錯差が大である時は、それが小なるものよりも合計點に及ぼす影響が大である。それ故に一般的學力を測定しようとする際には、先づ一般的學力の要素を決定し、次にその要素の重みを決定し、その次に重みに應じてその標準錯差を變更するに必要な乘除數を求め、この乘除數を粗點數に作用させて、然る後に合計點を求めなければならない。一般的讀方成績考査の規準を設定する場合でも理論上ではこれだけの手續をとらなければならないのである。

六　標準化した檢査の種類

多數の兒童生徒に對して、適切な試驗問題を課して、これを合理的に採點して得た代表値を有する檢査が、米國では各學科に亘つて多數に作られてゐるが、我が國に於ても若干のものが作られて居る。このやうに學力檢査の種類は多いが、使用目的に從つて、これを凡そ三種に大別し得る。一般的・診斷的及び練習的の三種卽ち是である。

〔一〕一般的學力測定　これは基礎的教材によつて兒童の凡その學力程度を決定するのを目的とする。米國のケレー、ルック及びターマンの三氏によつて作られたスタンフォード學力檢査の如きはこの一種で、讀方・算術・理科・歷史と

文學・語法・書取等の科目を含み、ピントナーの教育調査は讀方・算術・歴史・地理・語法等から成る。我が國に於て丸山は算術（計算と應用）・讀方（書取と語の意味）・國史・地理の四科目から成る一般學力檢査を構成して尋常五年の第三學期末と尋常六年の各月に對する代表値を設定したことがある。これ等は何れも各學科に亙る一般學力檢査であるが、これを讀方科に就いていへば、一般的讀方成績檢査を作ることが出來る。この際は例へば文字の書取・語の意味・語法・文章の大意把捉・讀書の速度等の如き各要素に亙る試驗問題を作つて、その合計點からして代表値を設定すればよい。

　この一般的讀方檢査は、校長なり視學にとつて有用な檢査である。校長はこれによつて各學級兒童の讀方進歩の程度を知ることが出來、視學は學校全體の進歩の程度を知り得るのである。これ監督官は各兒童が群團中でどういふ評點をとるかといふことよりも、全體としての學級成績の地位に興味を有してゐるからである。

　併しこれと共にまた教師もこの一般的學力檢査を用ひることが出來る。これを學年始と學年の終りに用ひるのである。今かりに尋常四年の受持となつたとする。教師は學年始めに於て受持兒童の讀方成績の大體を知つて、これによつて學級の讀方教授の方針を立てなければならぬであらう。教師はそれ故に尋常三年までの教材について、漢字の書取・語の意味・文の解釋・文章の大意把捉・讀書の速度等の綜合による一般的讀方檢査を施し、受持兒童の長所と短所とを明かにする必要がある。若し著しい短所があれば、次に述べる診斷檢査を施行して、これが原因を明かにしその救濟策を講じなければならぬ。學年の終りに一般的讀方檢査を再び施行して規準成績と比較する時は、第四學年受持の一ヶ年間に、どれだけ讀方成績を進歩させ得たか否かが數量的に分るのである。

〔二〕診斷的檢査　これはまた特殊的學力測定ともいふべく、前者の總合的なのに反して著しく分析的であり、前者の主として校長並びに視學に有用であつたのに對して、これは教師に有用な測定である。この檢査の目的とするところは兒童の特に優れてゐる點又は特に困難とする點を診斷するにある。兒童生徒が困難とするところが特殊的學力測定によつて分つたならば、更に進んで教師はその兒童に對して、口問口答を試み、或は課業後特別にその方面の學課作業を課してみるがよい。さうする時は判然とその困難とする原因が發見される筈である。原因が分ればこれに對して適當の處置を講ずればよい。

か〻る診斷は學力測定を行はないでも平素授業してゐれば、何れの教師にも分るところであるといふ人があるかも知れない。併し學力の優劣は教師の主觀的規準によつて判斷する時はとかく誤が多いもので、これは客觀的の規準によつて判定すべきものである。從つて標準成績のある檢査を施行してみなくては、兒童の學力の優劣は正確には分らないのである。

さて讀方成績の測定方面に就いては、文字の讀み・文字の書取・語の意味・文章の解釋・文章の大意把捉・語法・讀書の速度等の多方面があると思はれるが、こ〻では特殊學力測定の一例として、茗溪會研究部で標準化した書取檢査と、山本德行氏の默讀理解力檢査とを擧げてみよう。

〔1〕**茗溪會研究部の書取檢査**　これは兒童が日常生活に於て讀み書きするに必要な漢字の書取を考査する檢査を標準化したもので、選ばれた漢字は尋常小學國語讀本に提出されてゐる新出漢字一千三百五十六字である。その漢字を語句又は短文中に書き入れさすやうにして提出した。今その例を示せば次のやうである。

巻二

一、（オホ）キイ　ハナ。　二、（チヒ）サイ　マリ。　三九、（テン）マデ　アガレ。

巻三

一、（アメ）ガ　フル。　七一、キモノノ（ソデ）。

巻四

一、シゲツタ（モリ）。　一〇二、ヨク　キレル（カタナ）。

巻十二

一、アサヒガ　（ノボ）ル。　一一一、ワルイオコナヒヲ（セ）メテシカル。

巻二から巻十二まで、各巻に亙つて新出文字について、このやうな語句又は短文を以て提出した。

このやうな問題を作製して、昭和六年三月上旬から中旬にかけて四箇の尋常小學校第四學年並びに第六學年兒童について豫備的實驗を行つた。尋常科第四學年兒童には巻二から巻八までに提出されてゐる新出漢字の書取問題を課し、尋常科第六學年兒童には巻七から巻十二までの新出漢字を課した。檢査時間は制限しないで兒童全部が答へるまでの時間を與へた。

豫備的實驗の結果によつて、尋四・尋六の兩學年並びに男女別に各巻に於ける各漢字の正答者率を調査した。この檢査は時間制限でなかつたのであるから、各學年受驗者數を以て、正答者數を除し、これに一〇〇を乘じたものがここでいふ正答者率である。

各漢字の正答率を算出してから、尋常科第一・二・三學年用、尋常科第四・五・六學年用、並びに尋常科全學年用の三種の檢査系列を作成した。

尋常科第一・二・三學年用を「檢査一」と名づけ、これに對等の困難度を有する三類の檢査系列を作つて、これを甲類・乙類・丙類と名づけた。

尋常科第四・五・六學年用を「檢査二」と名づけ、これに對等の困難度を有する三類の檢査系列を作り、低學年用と同樣に甲類・乙類・丙類と名づけた。

尋常科全學年用は「尋常科全學年用」と名づけ、これに對等の困難度を有する甲・乙・丙・丁・戊の五類を作つた。

甲類一・乙類一・丙類一は尋常四年生に對する豫備的實驗の結果によつて、略對等の困難度を有する漢字から作成し、その範圍は卷二から卷七に及ぶこととした。かくて選ばれた漢字は各類とも二百二十二字であつた。

甲類二・乙類二・丙類二は尋常六年生に對する豫備的實驗の結果によつて、略對等の困難度を有する漢字から作成し、その範圍は卷七から卷十二に及ぶこととした。かくて選ばれた漢字は各類とも二百七十三字であつた。

尋常科全學年用は、卷二から卷六までの漢字については尋常四年生の豫備的實驗の結果により、卷七から卷十二までの漢字については尋常六年生の豫備的實驗の結果によつて、略對等の困難度を有する漢字を選んだ。卽ち甲・乙・丙・丁・戊の五類に亙つて何れも略對等の困難度を有する漢字を選び、その字數は各類とも百字であつた。

尋常科第一・二・三學年用、尋常科第四・五・六學年用、尋常科全學年の三種の檢査ともに、問題の排列に際しては、卷數の順序により、同一卷中では正答率の多いものから少いものへと排列した。

かくて作成された検査系列の一例を示せば次のやうである。

甲類一

一、オトナリ ノ （ヒト）。 二、（オホ）キイ ハナ。 三、タカイ （キ）。 四、（チヒ）サイ マリ。 五、カハ ノ （ミツ）

甲類二

一、ナガサガ 五（スジ）アルクギ。 二、ナツニナルト （ムシ）ガ タクサンデテクル。 三、タテト （ヨコ）。 四、フネガ （キシ）ニツク。

尋常科全學年用甲類

一、オトナリ ノ （ヒト）。 二、ヲトコ ノ （コ）。 三、エホン ヲ （ミ）ル。 四、（タカ）イヤネ。 五、ネムル（トキ）ガキタ。 六、（ナガ）イ タケザヲ。

検査系列を作つた上で、検査時間を決定した。この爲に用ひた検査系列はすべて甲類であつて、甲類一は尋常三年生に課し、甲類二と尋常科全學年は尋常六年生に課して、各兒童が何分間で自分の力で出來る限りの解答をなし終るかを一々記錄して、その後で適當な時間と思はれる長さを決定した。かくて各類一の検査時間は四十分間とし、各類二は五十分間とし、尋常科全學年の検査時間は二十分間とした。

標準成績設定の爲に選定した小學校は十一校であつて、甲類一・甲類二の標準成績はその中の九校の兒童の成績により、尋常科全學年用甲類の標準成績は十一校中の八校の兒童の成績によつて、検査は昭和七年三月に實施した。

いま標準成績即ち代表値を示せば次のやうである。

なほ檢査の種類に從ひ、各學年毎に正答數の頻數分配表を見る時は、檢査二は尋常四年・尋常五年ともに低い點の方に頻數が多い。尋常六年に於てもかゝる傾向がある。書取檢査としては、尋常四・五學年にはもつと易しい漢字をも加へて提出すべきであつた。この點に於ては尋常科全學年用の方が尋常四、五學年に對しては一層適してゐるといへる。

第三學期二月に於ける書取成績の代表値

檢査の種類	學年	男女計 受驗兒童數	男女計 平均點	男女計 標準錯差
甲類一	尋 1	947	10.4	3.9
	尋 2	1016	42.0	16.4
	尋 3	994	70.0	35.1
甲類二	尋 4	1050	37.4	29.1
	尋 5	1009	63.7	47.5
	尋 6	917	103.0	74.0
全學年用甲類	尋 1	791	3.4	1.0
	尋 2	802	10.9	3.9
	尋 3	790	17.8	7.9
	尋 4	819	26.4	12.4
	尋 5	802	34.6	18.3
	尋 6	636	51.0	26.5

前表の成績は第三學期二月に於ける代表値である。同一檢査間に於ては、相連れる學年成績間の差を十一ヶ月分して、低學年成績に累加することによつて、他の各月に對する代表値を算出することが出來る。そして次の公式を用ひればいつでも書取偏差値が求められる。

$$\text{書取偏差値}=\frac{\text{各自の書取成績點}-\text{平均點}}{\text{標準錯差}\times\frac{1}{10}}+50$$

書取偏差値五〇はその學年のその月の成績として普通であることを意味し、書取偏差値六五以上は書取成績の極めて優良であつて、三四以下は極めて劣等であることを意味する。

前述の代表値は甲類によつて得たものである。乙・丙等の他類は甲類と困難度を略等しくして作成したものであるが、果して甲類の規準が乙・丙等の他類に適用して差支ないであらうか。この各檢査類間の信頼性を、平均點と標準

錯差との類似度、點數相互間の相關關係によつて求めた。この實驗は尋常小學校兒童について試みたのであるが、その結果によれば嚴密な一致は得られなかつたけれども、大體に於てはよく一致してゐて、甲類で得た代表値はそのまま他の乙類丙類等に適用して差支ないやうである。

かくて我々は或學年で第一學期に甲類を使用して兒童の書取成績を測定し、第二學期にはその後の兒童の進步を知る爲に他の同程度の檢査を使用したい際には、乙類を用ひ、第三學期には丙類を用ひることが出來る。尋常科全學年用を用ひるとすれば、低學年の際は甲類又は乙類を用ひ、中學年では丙類又は丁類を高學年では戊類を使用し得る。

今かりに同一兒童に對して每學年書取檢査を課して、上表のやうな結果を示したとする。「は」兒童は尋常一年では

兒童名	智能偏差値	各學年書取偏差値					
		尋一	尋二	尋三	尋四	尋五	尋六
い	60	59	62	65	66	60	68
ろ	73	70	73	74	73	71	71
は	64	65	50	51	48	66	68
に	44	38	44	42	39	45	40
ほ	37	48	50	45	43	44	43
へ	64	52	60	52	54	49	46
と	53	55	58	51	52	54	58

智能に應じた成績を示してゐるが、次の三箇學年間では書取成績だけを見れば普通であるが、智能に比すれば著しく劣つて居る。それが尋常五年と六年に於て成績が復活して居る。「へ」兒童は智能偏差値は六四で優秀であるから、書取成績もこの程度に優秀であると期待されるのに、檢査成績によれば何れの學年に於ても豫期の成績を收めてゐない。一般に學年成績がかなり恒常的のものもあれば、或は昇り或は降つて動搖を示してゐるものもある。我々はかゝる測定を行ふことによつて始めて兒童の成績を正確に知ることが出來、そこに教授法の改善が企てられ、且は個々の兒童に對する指導が工夫されるのである。

〔2〕 **山本德行氏の默讀理解力檢査**　檢査材料選擇について、同氏は同好の人々に最初に尋常小學讀本卷五から高等小學讀本卷四までの十二冊中から、適當と思はれる文章百九十二を選んでもらひ、その中から同氏自らが適當と思はれるものを四十八選び、これを以て豫備檢査を施行し、その結果によつて二十の文を選ばれた。この二十の中十は要點把捉檢査に用ひ、十は大意把捉檢査に用ひた。

要點把捉檢査の實施に當つては先づ解答上の注意を與へ、次に練習を課し、次に本檢査を施されたもので、その檢査時間は五分間である。今十問題中若干例を示せば次のやうである。

〔一〕 一番はじめにつむのを一番茶といひます。その葉でこしらへる茶が一番よい茶になります。それから一月ほどたつてつむのを二番茶といひます。また三番茶、四番茶までもつむことがあります。

どんな茶が一番よい茶ですか。

1一番はじめにつんだ茶。2一月ほどたつてつんだ茶。3四番茶。4三番茶。5二番茶。

〔二〕 にいさんが歸つての話に、おぢさんの家は幸に無事であつたさうだ。火事はけいさつしよ前の米屋の土藏で止つたが、二棟の土藏の中一棟はとうとう燒落ちたといふことだ。郵便局や、けいさつしよも一時はあぶなくなつて大事な書類を取出した程であつたさうだ。

火事はどこで止まりましたか。

1郵便局。2米屋。3けいさつしよ。4米屋の土藏。5おぢさんのうち。

〔十〕 鳳凰室内部の装飾には、あらゆる技巧を施して、意匠は豐富に、手法は精美に、當時發達せる技術の奧妙を盡して、人を

してさながら極樂淨土に在るの感あらしむ。定朝の彫刻爲成の繪畫は建築の美と相待ちて三絶と稱すべく、唐土の美術の我が國民の嗜好に陶冶せられて漸く熟成せし國民的美術の精華は、全くこゝに聚められたりといふべし。

精華はこゝに聚められたりとはどこに聚められたのですか。

1鳳凰堂内部の裝飾に。2極樂淨土の中に。3定朝の彫刻の中に。4國民の嗜好の中に。5建築の美の中に。

大意把捉檢査でも最初に注意を與へ、練習を課して後十問題を提出し、その檢査時間を五分間とされた。その問題の若干例は次のやうである。

〔一〕汽車が走つてゐるときに、まどから外を見ますと、山も川も野原も林も後の方へとんで行くやうに見えます。田ではたらいてゐる人も、道を通つてゐる人も、馬も車も、今見えたかと思ふと、すぐに後になつてしまひます。

1山や川などのこと。2田ではたらいてゐる人のこと。3汽車が走るといふこと。4道を通つてゐる人のこと。5走つてゐる汽車のまどから外を見たときのこと。

〔五〕世界で一番高い家は、アメリカのニューヨークにあるウルワースビルディングである。地上から二百四十メートル餘も高くそびえ、それが五十五階になつてゐる。部屋の數は地下室を加へて千餘り、大ていは大會社大商店の事務室として使はれてゐる。又料理店や理髪店もあれば、りつぱな大水泳場もある。

1アメリカのニューヨークのこと。2世界で一番高い家のこと。3大會社や大商店の事務室のこと。4料理屋や理髪店のあること。5部屋の數がたくさんあること。

〔十〕手紙は目的や場合の異るに隨つて、其の文に精粗繁簡の趣を異にする必要がある。精密な説明を要する時には長きを厭はす委曲を盡して書くべく、父母に近況を知らせたり、友人の無聊を慰めたりする爲の書翰は格別として、普通の手紙は成るべく

簡潔を旨とするがよい。殊に急を要するものには、一言半句も贅言を挿まないやうにせねばならぬ。
1目的と場合とによつて手紙の長さを異にせよ。2手紙を送つて友人の無聊を慰めよ。3手紙は簡潔なのがよい。4父母に近況を知らす手紙のこと。5急を要する手紙の書き方。

山本氏がこの檢査系列作成上注意された點を述べれば次のやうである。

(一)尋常六年兒童に對する豫備的實驗の結果によつて、文章理解の困難度の順序に配列したこと。

(二)選擇肢の作製に注意し、提示した文章中に一度あらはれた語句を用ひたこと。

(三)文章理解の困難度に應じて採點上の重みをつけたこと。これは豫備的實驗によつて、正答人員の百分比を求め、これを標準錯差値に換算して點を算出したものである。かくて各問題に對する配當は二點から七點に及んで居り、全部正解する時は、兩檢査共に各四十五點をとることになる。

山本氏はこの檢査を尋常四・五・六學年の兒童に課して、その結果を下四分位數・中間數・上四分位數を以て示されてゐる。下四分位數とは全受驗兒童の最劣等のものから數へて四分の一の兒童の得た成績といふことであり、上四分位數は最劣等のものから數へて四分の三の位置にある兒童の成績といふことである。中間數は優と劣との中間の兒童の成績のことで、これは平均點と見てよい。

この結果による時は、下四分位數(劣)と上四分位數(優)との差が、尋常四年から五年を經て六年へと學年の進むにつれて増してゐる。また四分錯差を見てもその値が學年の進むと共に大きくなつて居る。それ等の結果からして理解力優劣の差は學年の進むと共に増大するといへる。山本氏は更にこの理解檢査の結果と、渡邊氏等の國民智能檢査の

要點把捉代表値

	尋四	尋五	尋六
下四分位數	3.0	6.1	8.2
中間數	6.3	11.1	14.1
上四分位數	10.6	15.4	19.4
四分錯差	4.0	4.7	5.6

大意把捉代表値

	尋四	尋五	尋六
下四分位數	3.7	7.6	8.7
中間數	7.2	11.2	13.4
上四分位數	11.3	16.4	18.7
四分錯差	3.8	4.4	5.0

結果との相關關係を求めて居られるが、それによると凡そ〇・六五の相關を示して居る。また要點把捉成績と大意把捉成績との相關は〇・六五（蓋然錯差〇・〇二五）であつたと報告されて居る。かゝる研究は正確な學力測定を行はなくては出來ないことである。

なほ文章鑑賞の標準化檢査について一言述べておきたい。若干の文學的價値の差等のある文章を、專門家及び學校教師等に提出してこれを評定せしめ、その結果によつて文學的順位を定めておく。これ等の文章を兒童に提示して、文學的價値に從つて順位をつけさせ、豫め定められた順位との差を求める。この順位の差の平均と標準錯差とを以て代表値を設定する。かうすることによつて個々の兒童の鑑賞眼の優劣を測定し得るのである。

〔三〕治療檢査　診斷檢査が教師のためのものであるとすれば、これは兒童生徒の爲の檢査である。この治療檢査の目的は檢査方法を用ひて、兒童をして自己指導の機會を與へようとするにある。診斷の結果兒童の短所が分つたならば、その短所とするところを矯正し治療する爲の練習問題檢査を課するのである。兒童は自分でその檢査に定められた指示に從つて、或は作業時間を測り、或は一定時間內に於ける作業成績を調べて、これを記錄し、或は自らの進步を知り、或は規準に比較して自分の讀書力の優劣を判定する。時としてはその結果を學習曲線に描かしめてもよい。

我が國に於てなほこの種の檢査の發達を見ないのは遺憾とするところである。標準化された檢査でなくても、教師

作製の客觀的檢査によつて練習しても有效であるが故に、兒童の成績劣等の原因が文字の讀み、漢字書取、文章の大意把捉等の如きものに存する場合は、その部分に對してこの客觀的檢査を與へて十分有效な練習をさせるがよい。若し兒童の智能が劣等なために讀方成績が劣るのであれば、その智能相應の材料の學習をもつて滿足しなければならぬ。又若し成績劣等の原因が眼や耳の缺陷にあるならば、これに對する適切な處置を講ずることが肝要である。

参考書

大伴茂「教育測定の原理と方法」昭和六年

岡部彌太郎「教育測定學」教育科學、第二十册、昭和八年

早瀬松藏「新應能教育の理論と實際」昭和八年

早瀬松藏「讀方考査に於ける評點のつけ方」教育心理研究、第四卷、昭和四年

山本德行「默讀理解力に關する研究」教育心理研究、第五卷、昭和五年

茗溪會研究部「國語書取成績考査規準」昭和九年

田中寛一「教育的測定學」大正十五年

田中寛一「教育的測定法」(師範大學講座)

田中寛一「選拔考査法」昭和六年

丸山良二「教育心理學」昭和八年

S. L. Pressey; Psychology and the New Education. 1933.

S. C. Garrison and K. C. Garrison: The Psychology of Elementary School Subjects. 1929.

國語科學講座
（第八回配本）

昭和九年四月十日印刷
昭和九年四月十五日發行

東京市神田區錦町一丁目十番地
編輯兼發行者 株式會社 明治書院
代表者 三樹退三

東京市神田區三崎町二丁目一番地
印刷者 細谷祐三

發行所 東京市神田錦町一丁目 株式會社 明治書院